DOCUMENT RESUME

ED 109 189 TM 004 639

AUTHOR Harmon, Judson A.; And Others
TITLE A Survey of Attitudes Toward Policies and Procedures in the Evaluation of Instruction.
PUB DATE [Apr 75]
NOTE 28p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, D. C., March 30-April 3, 1975); Not available in hard copy due to marginal legibility of original document

EDRS PRICE MF-$0.76 PLUS POSTAGE HC Not Available from EDRS
DESCRIPTORS *Educational Attitudes; Educational Objectives; Effective Teaching; *Evaluation Criteria; *Evaluation Methods; Observation; Statistical Analysis; *Surveys; *Teacher Evaluation; Teacher Rating

ABSTRACT

This paper describes the development and first administrations of a survey of attitudes toward policies and procedures in the evaluating of classroom teaching. The instrument was cooperatively developed by teachers, administrators, and consultants. It provides a collective picture of an ideal system of evaluation and, simultaneously, a picture of the adequacy of the present local system of evaluating instruction. Discrepancies between the real and the ideal or conflicts of attitudes between different respondent groups do occur and can be studied. Results for two school districts are presented. Implications for improvements in the survey and methodology are also presented. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

1975
AERA
Session # 22.20

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

A SURVEY OF ATTITUDES TOWARD POLICIES AND PROCEDURES IN THE EVALUATION OF INSTRUCTION

JUDSON A. HARMON
Metro Research
C.E.S.A. #19
Milwaukee, WI
53226

*BARRY FAGIN **
Admissions Office
Marian College
Fond du Lac, WI
54935

ROBERT B. INGLE
Ed. Psych. Dept.
U. of Wisc. - Milw.
Milwaukee, WI
53201

JAMES D. ROBINSON
Ed. Psych. Dept.
Marquette Univ.
Milwaukee, WI
53233

ED109189

TM 004 639 BEST COPY AVAILABLE

ABSTRACT

This paper describes the development and first administrations of a survey of attitudes toward a large number of policies and procedures in the evaluation of classroom teaching. The instrument, developed cooperatively by administrators, teachers, and consultants, asked how important each policy or procedure was in an ideal evaluation system and then how well it was being handled locally. Two Milwaukee area school districts were surveyed. Results showed some discrepancies between real and ideal systems and some group differences. Teachers definitely wanted to be involved at all stages. All stages were seen as being important. Methodological problems are discussed and some revisions suggested.

* * * * *

The evaluation of instruction has been a thorny issue since ancient times. Even Aristotle, in about 347 B.C., at Plato's Academy, was passed over when Plato died and the head position went to someone else. Aristotle, while departing for greener pastures, must have had some serious doubts about the fairness of the evaluation job done on him!

Of course we all reconginze that there are still many problems in the evaluation of teaching today. This is not only because the evaluator's task is a complicated and delicate one, but also because there are so many different

---

* Barry Fagin passed away on July 5, 1974. As with many other projects, he got this one moving and then handed it over to some interested friends to finish while he turned his attention to new problems.

opinions about what "good teaching" really is.

Another serious problem that evaluators of teaching may soon be required to deal with is new state legislation on the evaluation of educators. California's Stull Act is the first such legislation we have heard of. It requires each school district in California to adopt or develop its own system of evaluating certificated personnel. It is presented along with eight major presentations plus some thoughtful discussions in a conference report edited by N. L. Gage (1973).

There have been developed some workshop materials on the evaluation of teaching during the last three years. One of these, also from California (no coincidence), was developed at the Center for the Study of Evaluation at U.C.L.A. (Fink, et, al., 1973). It has recently been adopted for use in national workshops by the National Symposium for Professors of Educational Research. It covers: Teacher observations and ratings, Teacher skill testing, and Measurement of student growth as three general options in the evaluation or appraisal of teachers.

But this kind of workshop was probably forced by the Stull Act. Would it not be more advisable to find out what educators think is valuable before deciding upon one or more of the three options described above?

In the present study we tried to get a better idea of how school people perceive the evaluation process. We wanted to present a large number of specific policies and procedures and get judgements about how important or valuable they would be in an ideal system of evaluation. Then if the important policies were not being practiced in a district or building, the people who were responsbile could initiate those policies. Likewise, they could discontinue any policies or procedures seen as unimportant or trivial.

We were also interested in any policies or procedures for which there were disagreements between groups as to their importance. It seemed likely that administrators and teachers would see some items differently. The same for elementary versus secondary educators since their teaching tasks are rather different. Finally, teachers who had tenure might have different needs and values from those of non-tenured teachers.

Perhaps we looked with the most hope for those policies and procedures which were judged to be important across all groups of school people. These, we thought, might be the elements of a newer and better system of evaluation.

Procedures:

A survey instrument was developed in a cooperative project in which twelve public school districts in Milwaukee and Ozaukee counties particpated. This project was initiated by the late Barry Fagin. He and four school teachers met monthly during the 1972-73 school year and the other co-authors joined the project at various times after that.

During the Fall and Winter of 1973-74 the project was formalized and its efforts were focused upon constructing and piloting the survey. The survey was developed in a cooperative and interactive fashion. There were five preliminary forms of the survey that were developed, printed, studied, critiqued, defended, and revised by the participants before the present version was printed. The entire survey instrument is attached.

This survey is unique in three respects. First, it is the only survey we know of that assesses attitudes toward specific evaluation of teaching policies. Second, it is very comprehensive in scope; we asked the respondent to react to all the detailed issues that seemed to have some merit. Third, it allows for comparisons between how Important a policy or procedure is and how well it is being handled locally.

Its great comprehensiveness was seen as a necessary inconvenience, at least on the intial administrations. Unimportant items could be found empirically and then be removed for subsequent revisions. Seven operational phases or steps in the evaluation of classroom teaching were designated in the survey. They were arranged in a temporal order with three planning steps, three action steps, and one feedback step. These seven are listed below with an explanation for each one.

Somewhat briefer explanations were given to the respondents in the seven corresponding sections of the survey itself. The number of specific items for each of the seven sections is indicated in parentheses.

A. Selecting and Defining Educational Objectives. This section dealt with issues such as who should be responsible for specifying objectives or goals toward which teachers should aim and how should those objectives be determined? For example, one objective might be, "Each teacher will provide for individual differences." (14 items)

B. Determining Variables to be Employed in Evaluation. While Section A dealt with broad objectives or goals, Section B dealt with the manifestations or measurable overt behaviors which are associated with those goals. An example under the broad objective about providing for individual differences in A might be, "Each teacher should make available instructional materials which provide alternative reading difficulty levels." Both process and product evaluation were considered. (13 items)

C. Determining the Technique to be used in Gathering Evaluative Information. The main issue here was whether the evaluation of teaching should be based upon observational techniques or upon more objective techniques where some kind of survey or test instrument would be used. (13 items)

D. Guidelines for Observational Assessment. Assuming that some kind of observational procedures would be used, who should determine the procedures, who would do the observations, and what form should the observations take? (25 items)

E. Guidelines for Instrumental Assessment. Assuming that some kind of instruments (achievement tests, attitude surveys, etc.) were to be used, who should be responsible to select or develop those instruments and what form should they take? (12 items)

F. Deciding about Standards of Excellence in Teaching. Who should be responsible for setting up standards of excellence and how would these standards be applied in the whole evaluation system. (11 items)

G. How Should Evaluation Information be Shared? This section dealt with feedback sessions. It was assumed that the flow of information might go both ways between teacher and evaluator and not just the usual one-way flow from the evaluator to the teacher. (8 items)

The respondents were asked first to rate each item on the basis of its Importance on a 5-point scale from 1 ("Least Important") to 5 ("Most Important"). If they thought that the item "Should not even be considered," they were asked to signal this by marking an X instead of a number. Second, for the same item, they were asked to indicate how Effectively the policy or procedure had been handled in their own school's evaluation system. Again a 5-point scale was used from 1 ("Ineffective") to 5 ("Effective"). If the policy or procedure was not being practiced in their system, they were asked to signal this fact by marking an X again instead of a number.

A problem was anticipated by some of our participants. It seemed very likely that teachers would give high Importance ratings to some evaluation practices if they were to be used purely for the improvement of instruction but definitely not if those practices might subsequently be used against the teachers. Moreover, it would not be fair to advance an evaluation system for administrative decisions like promotion - non promotion, retention or dismissal, merit pay, work assignments, etc. when the evaluation system had been planned by teachers who thought it was going to be used for improving instruction. In order to overcome this problem, we had each respondent take the entire seven sections (96 items) two times. The first time they were to assume that their ratings would be used to shape an evaluation system purely for the improvement of instruction. The second time they were to assume that the eventual system would be used for the kinds of administrative personnel decisions mentioned above.

A final section of the survey, labeled "Part II", presented eight ways in which evaluations of teaching could be used, for example: to improve instruction or to reward superior teaching performances financially.

There was a total of 200 items. Of these, 192 had to do with evaluation policies and procedures and eight had to do with purposes of evaluation of teaching. The 192 figure represents 96 items taken under two assumptions. Since each item was rated with reference to both Importance and Effectiveness, each respondent made 400 separate ratings.

The respondents were from two suburban Milwaukee school districts. In one district, to which we will give the pseudonym, Birchwood, all teachers, administrators, supervisors, and specialists responded. In the second district, to which we will give the pseudonym, Pine Mountain, only the teachers and administrators of the elementary schools responded. The total N's were 134 and 74 respectively.

Since the response sheets were not optically scanable, we had all demographic and item response data keypunched. A computer program called COLFREQ (part of STAT-JOB) was used which reported the percentages of the group of respondents who gave each item ratings of 1, 2, 3, 4, 5, and X (keypunched as a blank). It also gave the means for each item.

It should be remembered that on any item some respondents marked an X for either Importance or Effectiveness or both. Accordingly, the mean ratings are based only on the responses of those people who made numerical responses, since X's were recorded as a categorically different response. As a result, there is no stable group of people whose responses are reflected in the item mean scores and it is not proper to use these means in any statistical test of the differences between items or of the differences between respondent groups. Nevertheless, mean ratings were calculated as a single, crude index of Importance or of Effectiveness. They are presented in categorical form with rather large class intervals. Percentage data were provided to the two districts but required too much space for this report.

Results:

1. Differences between ratings of importance under the two assumptions. While we had expected rather frequent differences between ratings under the two assumptions about how the eventual system of evaluation would be used, there were

almost no differences observed. Out of 96 items, only two items, A_6 and B_4, showed a noticable mean difference. Both of these were given higher Importance ratings under assumption #1 than assumption #2. Both items proposed including Parents in the process of evaluation. There is some sense to these results, but the mean differences were not very large and there were a large number of expected differences that did not show up. We concluded that any differences due to assumption #1 versus #2 were few enough and small enough to ignore and the results described below are based on responses under assumption #1 only (for improvement of instruction).

2. Isolation of crucial items and of rejected items. The item mean ratings for the Pine Mountain respondents are given in Table 1. The numbers in the cells refer to specific items within the seven sections, A - G, and Part II, which dealt with purposes of evaluating teaching. It is rather easy to spot items which were judged to be Important or Unimportant in each Section. Note that section G really had no items that were Unimportant or even of Moderate Importance. Note also that every section except Part II had at least four items with mean ratings of High Importance, (ie. over 4.0). Since items with a small number of respondents produced means that are somewhat suspect, those items are indicated with one or two asterisks. All but one of these (#D_{17}) had mean values below 3.0 which suggests that the mean values are not too misleading.

- - - - - - - - - - - - - - -

(INSERT TABLE 1 ABOUT HERE)

- - - - - - - - - - - - - - -

3. Isolation of items that are being well handled versus poorly handled locally. The item mean ratings on Effectiveness are given for the Pine Mountain district in Table 2. At first glance one is struck by the fact that no item received an Effectiveness rating of 4.01 or higher. Even with a progressive district there is room for improvement. An examination of the three Pine Mountain schools' individual means showed a sprinkling of values above the 4.01 level for one school or another, however, so the higher ratings are within reach.

TABLE 1:

MEAN RATINGS OF IMPORTANCE ACROSS ALL RESPONDENTS
IN PINE MOUNTAIN

SECTION	UNIMPORTANT		MODERATELY IMPORTANT		HIGHLY IMPORTANT	
	1.00-1.50	1.51-2.00	2.01-3.00	3.01-4.00	4.01-4.50	4.51-5.00
A		7*	5,6,8	1,2,4,9, 11,13	10,12,14	3
B	5*	4,6*	3*,7	12,13	2,8,9,10, 11	1
C	5**	3*,4*	6*,12	7	2,8,9,10, 11,13	1
D	5**,14**	4*,6*,13** 15*	3*,8,12*, 16,21	7,10,17** 22,24,25	2,9,11,18, 20,23	1,19
E	5**	3**,4**, 6*		7,10,11	2,9,12	1,8
F	5**	3*,4*	6*	7	2,9,11	1,8,10
G					1,3,6,8	2,4,5,7
PART II			2*,3	4,5,6,7,8		1

* and ** indicate items with 33% or more and 67% or more ratings of "X", respectively.

The second very noticable fact is that there are a very large number of items that are reported as not being used (note the one or two asterisks). We also can see that there are some items that had mean Effectiveness ratings of 3.01 or higher but were also marked with an X by many respondents (B_8, C_{13}, E_{11}, and F_{10}). Again, an inspection of the responses of the three individual schools accounts for the inconsistency. Usually two of the schools don't use the policy and the other one uses it rather Effectively; thus, we have many X's and a mean of over 3.01.

- - - - - - - - - - - - - - -

(INSERT TABLE 2 ABOUT HERE)

- - - - - - - - - - - - - - -

4. Isolation of items that are not being practiced locally. Table 3 presents these data for the survey, one section at a time, including the very important Part II items. If the item is on the right most people reported that it was NOT being practiced. It is not too surprising to see so many items being frequently reported as not being practiced since we tried to include innovative policies and procedures in the survey - especially in Section D, on the details of observational assessment techniques.

It is surprising and a bit disconcerting to see so many items in the middle ranges, from 21% to 80% X's. At first glance it would appear that the staff really doesn't know what policies are in effect and what ones are not in effect. However, there are several alternative explanations. Perhaps the item is being practiced in one building but not in others. Perhaps the item is being practiced on some respondents (e.g. recently hired teachers) but not on others. Finally, it is possible that the item was poorly written and was easily misunderstood by the respondents. This was a particularly fruitful part of the analysis for the administrators since it suggested that they might need to make it more clear to their teachers exactly what policies and practices are in force.

- - - - - - - - - - - - - - -

(INSERT TABLE 3 ABOUT HERE)

- - - - - - - - - - - - - - -

TABLE 2.

MEAN RATINGS OF EFFECTIVENESS ACROSS ALL RESPONDENTS

IN PIKE MOUNTAIN

SECTION	INEFFECTIVE 1.00-1.50	1.51-2.00	MODERATELY EFFECTIVE 2.01-3.00	3.01-4.00	VERY EFFECTIVE 4.01-5.00
A	7**	5* 6* 8*	9 11* 13* 14*	1 2 3 4 10 12	
B	3**	4*	1 5** 6* 7** 9* 10* 12 13**	2 8** 11	
C	5**	3** 4**	1* 6** 7** 8* 10* 12	2 9 11 13**	
D	3** 4** 5** 13** 14**	6** 15** 25**	1* 7* 8* 10* 12** 16** 17** 18 19 20 21** 22** 23* 24**	2 9 11	
E	3** 4** 5** 6**		1* 7* 10* 12*	2 8 9 11*	
F	3** 5**	4** 6**	1* 7** 8* 9* 11*	2 10*	
G			1 2 3 5 6 8*	4 7	

The * and ** marks indicate items with 33% and 67% or more ratings of "X", respectively.

TABLE 3:

PERCENTAGES OF "X" GIVEN TO ITEMS ON EFFECTIVENESS

SECTION	0-20%	21-40%	41-60%	61-80%	81-100%
A	1,2,3,4, 12	9,10,13 14	5,11	6,7,8	
B	2,11	1,10,12	9	3,4,6,7,8 13	5
C	2,11	1,9,12	8,10	6,7,13	3,4,5
D	2 9,19	11,18,20 23	1,8	6,7,10,12 13,21,25	3,4,5,14 15,16,17 22,24
E	2,8	1,9,11,12	7,10		3,4,5,6
F	2	10	1,8,11	6,7,9	3,4,5
G	1,2,3,4,5 7	6,8			
PART II	1,4,5,6	7,8	3	2	

5. Isolation of discrepancies between ratings of Importance and ratings of Effectiveness. Only the first section, "Selecting and Defining Educational Objectives," is presented here while all eight sections were reported in this fashion to the districts. Figure 1 shows the 14 items of Section A with regard to both their mean values of judged Importance (horizontal axis) and Effectiveness (vertical axis). One dashed line represents the median value on one axis and the other dashed line represents the median point on the other.

In the upper left quadrant thus formed, there are two items, 1 and 9, which we have somewhat arbitrarily identified as being well handled locally, but not very Important. In the lower right quadrant thus formed, there are two other items, 13 and 14, which we have identified as being quite Important but not very well handled locally.

There are other ways of pointing out such discrepancies but they are more complicated. For example, finding the items that depart the most from a least squares regression line is one good alternative.

- - - - - - - - - - - - - - -

(INSERT FIGURE 1 ABOUT HERE)

- - - - - - - - - - - - - - -

6. Recommendations of who should be involved in the seven phases of evaluating instruction. Teachers and Administrators were given high endorsements in all seven phases. On the other hand, Students, Community members, Parents, and School Board members were given such low ratings we could conclude that they were rejected as potential participants in all phases. The teachers seemed to be more interested in self evaluation than having their colleagues observe and evaluate them. Finally, Educational Experts were consistently given medium ratings as potential participants in the evaluation process.

7. Ratings of purposes of evaluation. The most Important purpose of evaluating teaching was to help teachers improve their teaching effectiveness. This can be seen in Table 1. However, it is also apparent that the other uses of evaluating teaching are not so trivial or aversive as we had anticipated.

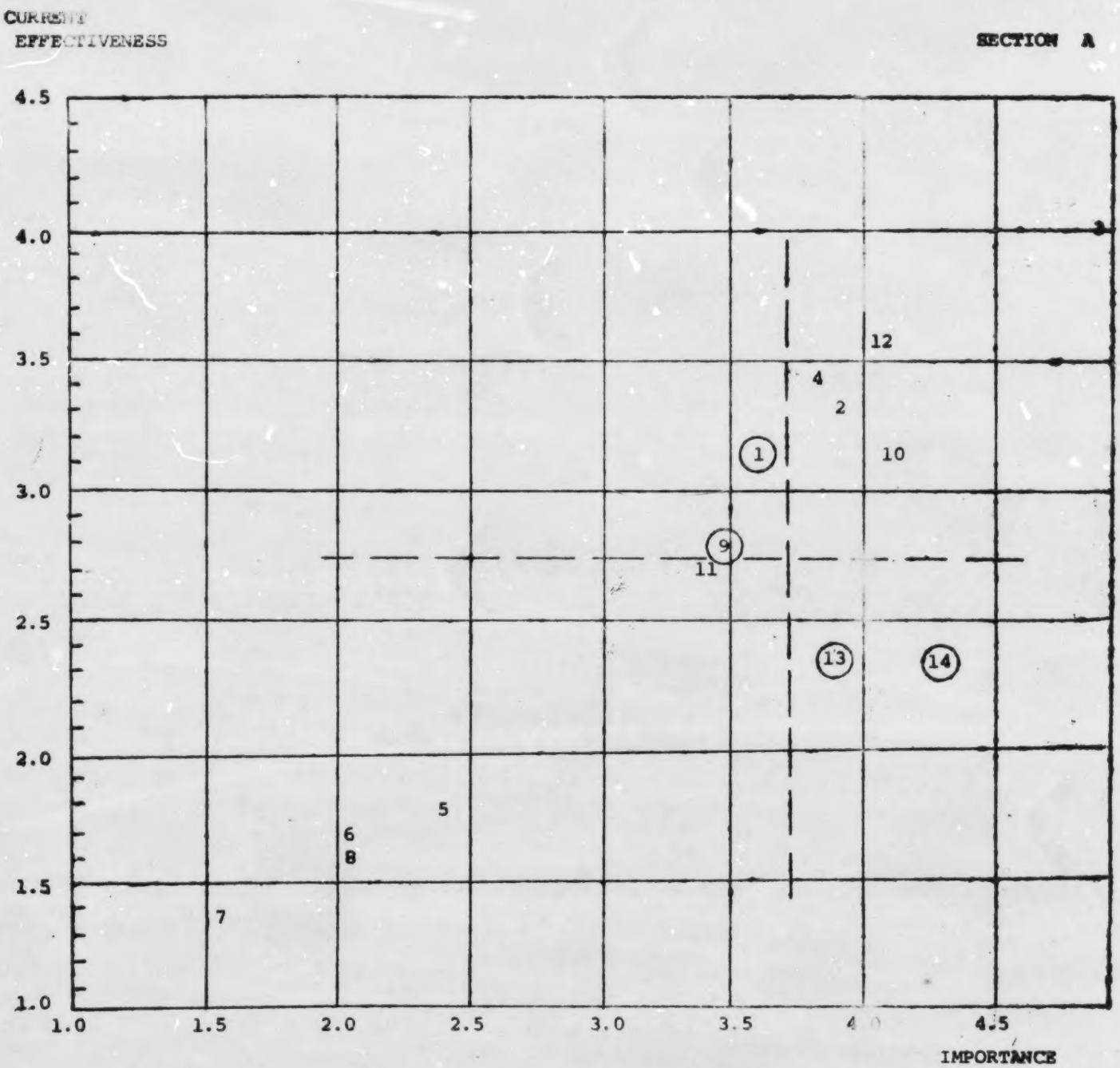

FIGURE 1. ITEM MEAN RATINGS OF IMPORTANCE AND EFFECTIVENE
FOR THE SECTION ON SELECTING AND DEFINING EDUCATIONAL OBJECT ES.

8. Differences between groups. All of the following results were observed in the analysis of Birchwood data. There were too few administrators who responded to the survey to compare them, as a group, with all the teacher respondents. However, several other interesting group differences did occur. All of these had to do with who should be involved as participants in the various phases of evaluation. For example, the Secondary educators were more willing than the Elementary educators to involve Board Members or Students in the various phases of the process. But they were less willing than Elementary educators to involve Community members and Educational Experts. In the comparisons between Tenured and Non-Tenured teachers, the Tenured teachers gave somewhat higher Importance ratings to items A_6, C_7, D_7, F_6, and F_7. Three of these items indicated a greater importance placed by Tenured teachers on Educational Experts participating in various phases and one item each for Parents and for Board Members participating.

Discussion

This project provided some methodological insights as well as the factual findings outlined above. First, a small pilot study should have been conducted to see if there were differences in Importance ratings under the two assumptions. There were very few differences and we effectively wasted some of the time of a large number of people.

Second, a rather careful set of directions was developed for use in the Birchwood survey which seemed to help the respondents get interested in and respond to the survey better than was the case at Pine Mountain.

Third, the respondent's task was a rather complicated one, rating both Importance and Effectiveness and marking either an X or a rating from 1 to 5 on both issues. This probably contributed to measurement error because of some respondents' confusion and irritation. It might be better to eliminate the "X" response entirely.

Fourth, even when only one assumption is used, the survey is quite time-consuming and should be shortened. A revised form will be ready in late April which will have only 75 items. It should be much easier for the respondents to take.

Fifth, while this survey does help to identify areas of discrepancy between Importance and Effectiveness, it does not help the school people in determining exactly what is wrong with those discrepant policies or procedures. We would encourage the school administrations to devote some time in committee with teachers or in one-to-one interviews exploring the specific parts of the system that are not working satisfactorily after they have been identified.

Finally, this survey or its successor should be studied and analysed with regard to several obvious technical points. These would include factorial validity, subtest homogeneity, item-scale correlations, subtest orthogonality, test-retest reliability, and the like.

REFERENCES

Fink, Arlene, Klein, S.P., and Burry, J., Teacher appraisal for improvement. Los Angeles: Center for the Study of Evaluation, 1973.

Gage, N.L., Ed., Mandated evaluation of educators: A conference on California's Stull Act, Washington, D.C.: Educational Resources Division, Capitol Publications, Inc., 1973.

* * * * * * * * * * * * * *

A PAPER PRESENTED AT A.E.R.A. ON WEDNESDAY, APRIL 2, 1975

IN WASHINGTON, D.C.

* * * * * * * * * * * * * *

RESPONSE INSTRUCTIONS

ITEM CONTENT

Part I consists of a number of items each of which describes either a procedure or a policy which can be employed in the development or operation of an evaluation system.

Examples

a. A Policy Employed in Development
(Item) Involve teachers in the selecting and defining of the educational objectives upon which teacher evaluations are to be based.

b. A Procedure of Operation
(Item) Have each observational visit long enough to allow the observer to get a clear picture of what the teacher is doing.

RATING TASK

You will be asked to rate each item in two different ways. First, you are to rate the item in terms of how important you believe it is to include the policy or procedure described in the item in any new evaluation system we might develop. Secondly, you are being asked to rate our present evaluation system in terms of how effectively it handles the policy or procedure described in the item.

SPECIFIC RATING DIRECTIONS

1. Read each item carefully.
2. Decide in your own mind how important the procedure or policy described in the item is.
3. Look at the Importance scale on the top of the response sheet to find the rating (number) which best describes your feelings about the item's importance.
4. Place this number in the square ☐ which appears to the left of the item. Use "X" if item should not be considered.
5. Now think about how effectively our present system deals with the policy or procedure.
6. Look at the Effectiveness scale on the top of the response sheet and find the rating (number) which best represents your feelings.
7. Place that number in the circle ◯ which appears to the right of the item you are rating. Mark an "X" in the circle if the policy or procedure is not practiced.
8. Go to each successive item until you have completed the questionnaire.
9. When you have finished, check your response sheet to make sure you have rated every item twice. See if you have placed a number in both the square and the circle for all the items.

CAUTION

Before you start responding we would like to point out three things to you. First, your Importance ratings will be used to determine what policies and procedures are the most crucial to include in any new evaluation system. Second, your feelings about Effectiveness will identify those procedures and policies with which our present system has trouble. This will tell us what types of changes are needed in any modifications for improvement of our present system.

Third, a technical note: If you place a lot of items in just one or two categories (for example, rate most items 4 or 5] you will, in effect, be diluting your own vote on making these decisions. One way to avoid this and to make sure your responses will have a maximum influence on these change decisions is to try to place at least four items from the entire questionnaire in each of the five rating categories, both for Importance and for Effectiveness.

When responding to all parts of the questionnaire assume that you will have approximately your present amount of influence upon such factors in the classroom as class size, student arrangements, materials, other resources, etc. that influence student and teacher behaviors.

PART I
Section 1

To the Respondent:

The responses you make in this section could be used to form an evaluation system which would provide information solely for the use of individual teachers in improving instruction. With this assumption in mind start with item A_1 and give your ratings of Importance and Effectiveness.

A second section will ask you to respond assuming that the evaluation system would be used to provide information to administrators. This information could be used by them for making such decisions as Promotion, Retention-Dismissal, Tenure-Non Tenure, Merit Considerations and Job Assignments.

Remember, when responding to this section assume that you will have approximately your present amount of influence upon such factors in the classroom as class size, student arrangements, materials, other resources, etc. that influence student and teacher behaviors.

A. Selecting and Defining Educational Objectives

The term educational objectives includes both instructional and professional objectives. Instructional objectives can be defined as desirable changes in students' knowledge and behavior. Professional objectives include such things as teachers' self improvement, assisting other teachers in increasing their effectiveness or carrying out established administrative policies.

Educational objectives could be employed as the basis for deciding what will be observed or measured in order to evaluate teaching. With this in mind, indicate on the response sheet how important it is to and how effectively has our present system been able to:

1. Develop a written list of specifically defined objectives toward which every teacher is expected to work.
2. Develop additional objectives which are more specific to grade level and subject or departmental areas.

3-9. Include representatives of the following groups in establishing objectives.

3. teachers
4. administrators
5. students
6. parents
7. community members
8. school board members
9. educational experts

10. Provide the teacher with a written list of general objectives for all teachers and specific objectives for teachers at a given grade level or in a particular subject area.
11. Give different weights to different objectives (i.e. assign priority to them) on the basis of a rating inventory (like the one you are taking) given to all teachers.
12. Allow teachers to add to the general list of objectives any they personally feel are important.
13. Provide an in-service session for clarifying the general list of objectives.
14. Provide a regular and systematic way to review and revise the list of objectives.

B. Determining Variables to be Employed in Evaluation

There are basically only two means of assessing a teacher's ability to achieve a given educational objective. It can be determined either by a process evaluation or a product evaluation. Process evaluation focuses on what a teacher does in order to provide good instruction while product evaluation focuses on those student behaviors which logically should be affected as a result of good instruction.

It is possible and perhaps desirable to employ a combination of both forms of evaluation. With this in mind, indicate how important it is and how effectively has our present system been able to:

1-7. Include representatives from the following groups in selecting the variables to be used in assessing teachers' ability:
   1. teachers
   2. administrators
   3. students
   4. parents
   5. community members
   6. school board members
   7. educational experts

8. Provide an in-service meeting at which each of the variables selected will be clarified and justified. (Objective by objective.)
9. Allow individual teachers to add to the variables selected any additional ones which they feel would also be appropriate in assessing their ability.
10. Provide some systematic way of reviewing and revising the list of variables used in evaluation.
11. Evaluate teaching by the use of process variables (teacher actions).
12. Evaluate teaching by the use of product variables (consequent student actions).
13. Let each teacher decide whether his/her teaching is to be evaluated by Process or Product Variables or both and try the choice for at least one year.

C. Determining the Technique to be Employed in Gathering Evaluative Information

Basically there are only two means available for gathering data on the variables which are selected. They are the observational and instrumental techniques.

Observational techniques are measurements obtained from an observer (ratings) who has watched or listened to a teacher at work (for process variables) or students' performance (for product variables).

The instrumental technique involves measurements obtained from student tests (achievement, attitude etc.), surveys, questionnaires, or some other sort of instrument which samples directly a teacher's behaviors (process) or the students' behaviors (product).

With these definitions in mind indicate how important it is and how effectively has our present system been able to:

1-7. Include representatives of the following groups in deciding whether observations and/or instruments should be used in measuring a given variable:
   1. teachers
   2. administrators
   3. students
   4. parents
   5. community members
   6. school board members
   7. educational experts

8. Provide an in-service session for examining and understanding materials and procedures to be used in measuring.
9. Allow teachers to go beyond the limits of the particular technique chosen and add measures they believe to be important and wish to use.
10. Provide a time and a system for reviewing and revising the decisions about methods for data gathering.
11. Evaluate teaching by means of observational techniques.
12. Evaluate teaching by means of instrumental techniques.
13. Let each teacher decide whether his/her teaching is to be evaluated by observational or instrumental techniques or both and try the choice for at least one year.

D. Guideline for Observational Assessment

Assume the observational technique has been chosen to measure some variable. How important is it and how effectively has our present system been able to:

1-7. Include representatives from the following groups in determining who will observe the teachers' or students' behaviors chosen as variables, outline the procedures to be followed by the observer, and develop or select the materials to be employed:
   1. teachers
   2. administrators
   3. students
   4. parents
   5. community members
   6. school board members
   7. educational experts

8-17. Include members of the following groups as observers:
   8. central office administrative staff
   9. principals
   10. peers
   11. self
   12. students
   13. parents
   14. community members
   15. school board members
   16. educataional experts
   17. others (please specify on response sheet)

18. Provide for sufficient number of formal observations per term to obtain a precise measure on each variable.
19. Provide for a sufficient amount of time on each formal observation to allow the observer to get a clear picture of what the teacher or student is doing.
20. Provide for sampling (formally observing) the same behavior during different times of the day.
21. Allow each teacher to schedule the day and time of the majority of the formal observations.
22. Provide a training program for the observers.
23. Provide the observer with enough background information about the students (skill level, etc.) or teacher (objectives, lesson plans, etc.) to be observed for him or her to understand what is going on.
24. Provide an in-service session for trying out and getting used to new modes of observation (e.g. video tapes, etc.).
25. Provide a regular time and process for reviewing and revising the observational programs (including who observes, materials used, procedures followed, and reports given).

E. Guidelines for Instrumental Assessment

Assuming some instruments (achievement tests, attitude tests, surveys etc) are being employed to measure one or more variables, how important is it and how effectively has our present system been able to:

1-7. Include representatives from the following groups in selecting and/or developing the instruments to be used in measuring variables:
   1. teachers
   2. administrators
   3. students
   4. parents
   5. community members
   6. school board members
   7. educational experts

8. Make an effort to be sure that the tests to be used are valid and reliable.
9. Provide teachers with written examples of the items contained in the instruments to be used in assessment.
10. Allow individual teachers to add any items or subjects which they think are important to obtaining an accurate measurement.
11. Schedule an in-service session at which the tests, surveys, etc. would be explained, demonstrated, and defended.
12. Provide a regular time and system for reviewing and revising the instruments and their items.

F. Deciding About Standards of Excellence in Teaching

In any evaluation of teaching system it is possible to establish minimum standards which must be met and some graduations above the minimum standard so that all teachers can be adequately and fairly assessed.

With this in mind indicate how important it is and how effectively our present system has been able to:

1-7. Include representatives of the following groups in determining what are minimum standards and what graduations should be determined above the minimum standards:
   1. teachers
   2. administrators
   3. students
   4. parents
   5. community members
   6. school board members
   7. educational experts

8. Provide a written statement about standards to teachers.
9. Provide an in-service session for explaining and justifying all the standards.
10. Provide for adjusting the standards for teachers with unusual class size, work loads, poorer or low achievers, etc.
11. Provide for a regular system for evaluating and revising standards.

G. How Should Evaluation Information Be Shared

This section refers to procedures which could be followed in order to provide information to the teachers regarding the results of his/her evaluation.

With this in mind tell how important is it and how effectively has the present system been able to:

1. Provide for frequent evaluation sessions during the year.
2. Schedule feedback sessions as soon as possible after each of these evaluation sessions (while things are still fresh in people's minds).
3. Have a written report to accompany the feedback session (so that forgetting of verbal communications is not a problem).
4. Allow or even encourage the individual teacher to provide additional evaluative information he or she has acquired independently to corroborate or refute the written report.
5. Focus part of the session upon possible changes the teacher could make (alteration of style or approach, use of different materials, etc.) in order to improve or vary instruction.
6. Provide the needed resource materials and ideas to help a teacher who wishes to try some innovative approach to improve his or her ability.
7. Provide a relaxed atmosphere that encourages teachers to engage in self evaluation and to take an active part in determining the direction of the feedback session.
8. Provide a system by which the feedback procedures can be changed when necessary.

PART I
Section 2

To the Respondent:

The responses you make in this section could be used to form an evaluation system which will provide information to administrators. This information could be used by them for making such decisions as Promotion, Retention-Dismissal, Tenure-Nontenure, Merit considerations and Job Assignments.

When responding to this section assume that you will have approximately your present amount of influence upon such factors in the classroom as class size, student arrangements, materials, other resources, etc. that influence student and teacher behaviors.

We are asking you to rate the same items as before -- only with the difference in how your ratings may be used. So go back to item A_1 and answer each item with these new uses in mind.

After you have finished Part I, go on to Part II on the following page. It is short and very important.

PART II

To the Respondent:

The items in this final part are not really policies or procedures of an evaluation system. They are, instead, ways to use evaluation information. Tell how important each item is and how effectively our present system has used evaluation of teaching information to:

1. Help teachers improve their teaching effectiveness.
2. Reward superior performance financially.
3. Determine tenure.
4. Make decisions regarding staff assignments.
5. Make decisions regarding retention and dismissal.
6. Protect the teacher in legal matters.
7. Protect the school district in legal matters.
8. Validate the instruments and systems used in personnel selection and assignment.